Automatica 48 (2012) 1759–1767

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Multivariable Newton-based extremum seeking☆

Azad Ghaffari [a], Miroslav Krstić [b,1], Dragan Nešić [c]

[a] Joint-Doctoral Programs (Aerospace and Mechanical), San Diego State University and University of California at San Diego, La Jolla, CA 92093-0411, USA

[b] Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0411, USA

[c] Department of Electrical and Electronic Engineering Department, University of Melbourne, VIC 3010, Australia

ARTICLE INFO

Article history:
Received 11 May 2011
Accepted 21 January 2012
Available online 26 June 2012

Keywords:
Newton-based extremum seeking
Averaging
Singular perturbation
Stability

ABSTRACT

We present a Newton-based extremum seeking algorithm for the multivariable case. The design extends the recent Newton-based extremum seeking algorithms for the scalar case and introduces a dynamic estimator of the inverse of the Hessian matrix that removes the difficulty with the possible singularity of a possible direct estimate of the Hessian matrix. The estimator of the inverse of the Hessian has the form of a differential Riccati equation. We prove local stability of the new algorithm for general nonlinear dynamic systems using averaging and singular perturbations. In comparison with the standard gradient-based multivariable extremum seeking, the proposed algorithm removes the dependence of the convergence rate on the unknown Hessian matrix and makes the convergence rate, of both the parameter estimates and of the estimates of the Hessian inverse, user-assignable. In particular, the new algorithm allows all the parameters to converge with the same speed, yielding straight trajectories to the extremum even with maps that have highly elongated level sets, in contrast to curved "steepest descent" trajectories of the gradient algorithm. Simulation results show the advantage of the proposed approach over gradient-based extremum seeking, by assigning equal, desired convergence rates to all the parameters using Newton's approach.

1. Introduction

Motivation. Dramatic advances have occurred over the past decade both in the theory (Ariyur & Krstić, 2002, 2003; Choi, Krstić, Ariyur, & Lee, 2002; Krstić & Wang, 2000; Rotea, 2000; Stankovic, Johansson, & Stipanovic, 2010; Stankovic & Stipanovic, 2010; Tan, Nešić, & Mareels, 2006; Tee1 & Popovic, 2001; Wang & Krstić, 2000) and in applications (Banaszuk, Ariyur, Krstić, & Jacobson, 2004; Becker, King, Petz, & Nitsche, 2007; Carnevale et al., 2009; Cochran, Kanso, Kelly, Xiong, & Krstić, 2009; Cochran & Krstić, 2009; Guay, Perrier, & Dochain, 2005; Killingsworth, Aceves, Flowers, Espinosa-Loza, & Krstić, M, 2009; Luo & Schuster, 2009; Wang, Yeung, & Krstić, 2000; Zhang, Arnold, Ghods, Siranosian, & Krstić, 2007) of extremum seeking control. All these references employ gradient-based extremum seeking.

A Newton-based extremum seeking algorithm was introduced in Moase, Manzie, and Brear (2010) where, for the single-input case, an estimate of the second derivative of the map was employed in a Newton-like continuous-time algorithm. A generalization, employing a different approach than in Moase et al. (2010), was presented in Nešić, Tan, Moase, and Manzie (2010), where a methodology for generating estimates of higher-order derivatives of the unknown single-input map was introduced, for emulating more general continuous-time optimization algorithms, with a Newton algorithm being a special case.

The key distinction of the Newton algorithm relative to the gradient algorithm is that, while the convergence of the gradient algorithm is dictated by the second derivative (Hessian) of the map, the convergence of the Newton algorithm is independent of the Hessian and can be arbitrarily assigned. This is particularly important in non-model based algorithms, like extremum seeking, where the Hessian is unknown.

The power of the Newton algorithm is particularly evident in multi-input optimization problems. With the Hessian being a matrix in this case, and with it being typically very different from the identity matrix, the gradient algorithm typically results in different elements of the input vector converging at vastly different speeds. This problem is inherent to gradient-based schemes. To rectify it one would need to modify the algorithm using the inverse of the Hessian matrix which is not available as the model of the system is assumed to be unavailable. On the other hand, the Newton algorithm, if equipped with a convergent estimator of the Hessian matrix, achieves convergence of all the elements of the input vector at the same, or at arbitrarily assignable, rates.

☆ This research was supported by the Australian Research Council under the Discovery Grants scheme. The material in this paper was presented at the 50th IEEE Conference on Decision and Control (CDC 2011), December 12–15, 2011, Orlando, FL, USA. This paper was recommended for publication in revised form by Associate Editor Warren E. Dixon under the direction of Editor Andrew R. Teel.

E-mail addresses: aghaffar@ucsd.edu (A. Ghaffari), krstic@ucsd.edu (M. Krstić), d.nesic@ee.unimelb.edu.au (D. Nešić).

[1] Tel.: +1 858 822 1374; fax: +1 858 822 3107.

0005-1098/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2012.05.059

Results of the paper. In this paper we present a multivariable Newton-based extremum seeking algorithm, which yields arbitrarily assignable convergence rates for each of the elements of the input vector. We generate the estimate of the Hessian matrix by generalizing the idea proposed in Nešić et al. (2010) for the scalar case.

Generating an estimate of the Hessian matrix in non-model based optimization is not the only challenge. The other challenge is that the Newton algorithm requires an inverse of the Hessian matrix. The estimate of this matrix, as it evolves in continuous time, need not necessarily remain invertible. We tackle this challenge by employing a dynamic system for generating the inverse asymptotically. This dynamic system is a filter in the form of a Riccati differential equation. When fed with a positive/negative-definite estimate of the Hessian matrix over a longer period of time, this filter converges to a positive/negative-definite inverse of the Hessian matrix. Hence, after a transient, our non-model based algorithm behaves (on average) as a model-based Newton algorithm.

While the basic idea of our algorithm is developed for static maps, we provide the analysis of convergence when the algorithm is applied to general nonlinear systems, as in Krstić and Wang (2000). We apply classical averaging and singular perturbation methods, so our stability result is local—the parameter estimates start not too far from the true parameters and the estimate of the Hessian matrix starts not too far from the true Hessian matrix. It is also possible to prove non-local stability of the proposed scheme in a similar manner as Tan et al. (2006) where a gradient-based algorithm was investigated.

The continuous-time Newton algorithm that we propose is novel, to our knowledge, even in the case when the cost function being optimized is known. The state-of-the-art continuous-time Newton algorithm in Airapetyan (1999) employs a Lyapunov differential equation for estimating the inverse of the Hessian matrix, see (3.2) in Airapetyan (1999). The convergence of this estimator is actually governed by the Hessian matrix itself. This means that the algorithm in Airapetyan (1999) removes the difficulty with inverting the estimate of the Hessian, but does not achieve independence of the convergence rate from the Hessian. In contrast, our algorithm's convergence rate is independent from the Hessian and is user-assignable.

Organization. We state the problem and review the gradient-based extremum seeking algorithm for a static map in Section 2. Section 3 presents our Newton-based scheme for the static map. In this section we explain how we generate the estimate of the Hessian matrix and the estimate of its inverse. A generalization of the Newton-based scheme to dynamic plants is introduced in Section 4. The main stability result is stated in Section 5. Stability analysis based on the averaging and singular perturbation methods is presented, respectively, in Sections 6 and 7. Section 8 presents an illustrative example to highlight the difference between the proposed scheme and the standard gradient-based extremum seeking. (See Fig. 1.)

2. Review of the gradient algorithm for static map

Consider a convex static map

$$y = Q(\theta), \quad \theta = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_n]^T, \tag{1}$$

with a local maximum at θ^*. The cost function is not known in (1), but we can measure y and we can manipulate θ. The gradient-based extremum seeking scheme for this multivariable static map is shown in Fig. 2, where K is a positive diagonal matrix, and the perturbation signals are defined as

$$S(t) = [a_1 \sin(\omega_1 t) \quad \cdots \quad a_n \sin(\omega_n t)]^T, \tag{2}$$

$$M(t) = \left[\frac{2}{a_1}\sin(\omega_1 t) \quad \cdots \quad \frac{2}{a_n}\sin(\omega_n t)\right]^T, \tag{3}$$

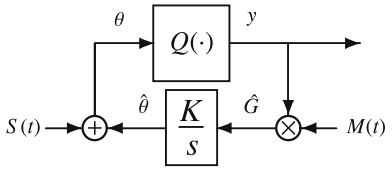

Fig. 1. Gradient-based extremum seeking for a static map.

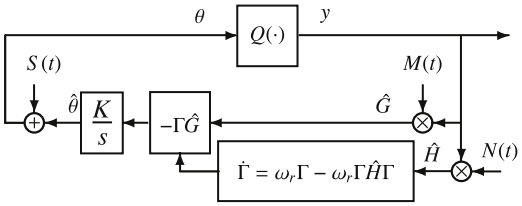

Fig. 2. Newton-based extremum seeking for a static map.

where ω_i/ω_j are rational and a_i are real numbers, with the frequencies chosen such that $\omega_i \neq \omega_j$ and $\omega_i + \omega_j \neq \omega_k$ for distinct i, j, and k.

Remark 1. A gradient-based extremum seeking for the static map (1) is given by $\dot{\hat{\theta}} = KM(t)y$, $\theta = \hat{\theta} + S(t)$. In the parameter error variable $\tilde{\theta} = \hat{\theta} - \theta^*$, the closed-loop system in Fig. 2 is given by $\dot{\tilde{\theta}} = KM(t)Q(\theta^* + S(t) + \tilde{\theta})$. The basic idea of the scheme, as well as of the choice of the perturbation signals, is understood by noting that, for the case of a quadratic map, $Q(\theta) = Q^* + \frac{1}{2}(\theta - \theta^*)^T H(\theta - \theta^*)$, the averaged system is given by

$$\dot{\bar{\tilde{\theta}}} = KH\bar{\tilde{\theta}}, \tag{4}$$

where H is the Hessian of the static map, and it is negative definite. This observation reveals two things: (i) the gradient-based extremum seeking algorithm is locally convergent, and (ii) the convergence rate is governed by the unknown Hessian matrix H. One of the features of the Newton algorithm presented in the next section is to eliminate the dependence of the convergence rate on the unknown H.

3. Newton algorithm for static map

The Newton-based extremum seeking algorithm for a static map is shown in Fig. 2, where ω_r is a positive real number. There are two vital parts in the Newton-based algorithm: the perturbation matrix $N(t)$, which generates an estimate of the Hessian, and the Riccati equation, which generates an estimate of the inverse of the Hessian matrix, even when the estimate of the Hessian is singular.

The idea for producing the estimate of the Hessian matrix $H := \partial^2 Q(\theta^*)/\partial\theta^2$ is borrowed from the scalar design in Nešić et al. (2010). Referring to the Taylor series expansion of the cost function around the peak, we have

$$\begin{aligned} y &= Q(\theta^* + \tilde{\theta} + S(t)) \\ &= Q(\theta^*) + \frac{1}{2}(\tilde{\theta} + S(t))^T H(\tilde{\theta} + S(t)) + R(\tilde{\theta} + S(t)), \end{aligned} \tag{5}$$

where $\partial Q(\theta^*)/\partial\theta = 0$ and $R(\tilde{\theta} + S(t))$ stands for higher order terms in $\tilde{\theta} + S(t)$. The product of $N(t)$ and y needs to generate an estimate of the Hessian in an average sense. We show that by an appropriate selection of matrix $N(t)$, the average value of $\dot{\hat{H}} = N(t)y$ over the period Π, which is related to ω_i's (see (10)), is an estimate of the Hessian. We start with

$$\frac{1}{\Pi}\int_0^{\Pi} N(\sigma)y\mathrm{d}\sigma$$
$$= I + J + \bar{H} + \frac{1}{\Pi}\int_0^{\Pi} R(\tilde{\theta} + S(\sigma))N(\sigma)\mathrm{d}\sigma, \tag{6}$$
$$I := \frac{1}{\Pi}\int_0^{\Pi}\left(Q(\theta^*) + \frac{1}{2}\tilde{\theta}^T H\tilde{\theta}\right)N(\sigma)\mathrm{d}\sigma, \tag{7}$$
$$J := \frac{1}{\Pi}\int_0^{\Pi}\tilde{\theta}^T HS(\sigma)N(\sigma)\mathrm{d}\sigma, \tag{8}$$
$$\bar{H} := \frac{1}{\Pi}\int_0^{\Pi}\frac{1}{2}S(\sigma)^T HS(\sigma)N(\sigma)\mathrm{d}\sigma$$
$$= \frac{1}{\Pi}\int_0^{\Pi}\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}H_{i,j}\sin(\omega_i\sigma)\sin(\omega_j\sigma)N(\sigma)\mathrm{d}\sigma. \tag{9}$$

By taking Π as

$$\Pi = 2\pi \times \mathrm{LCM}\left\{\frac{1}{\omega_i}\right\}, \quad i \in \{1, 2, \cdots, n\}, \tag{10}$$

where LCM stands for the least common multiple, we have $I = 0$ if N has zero average over Π. Also, taking N such that

$$\frac{1}{\Pi}\int_0^{\Pi}\sin(\omega_i\sigma)N_{j,k}(\sigma)\mathrm{d}\sigma = 0, \tag{11}$$

holds for all $i, j,$ and $k \in \{1, 2, \ldots, n\}$, makes the integral J equal to zero. Furthermore, $\bar{H}$ is equal to H if we choose N such that

$$\frac{1}{\Pi}\int_0^{\Pi}\sin^2(\omega_i\sigma)N_{i,i}(\sigma)\mathrm{d}\sigma \neq 0 \tag{12}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin(\omega_i\sigma)\sin(\omega_j\sigma)N_{i,j}(\sigma)\mathrm{d}\sigma \neq 0 \tag{13}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin^2(\omega_i\sigma)N_{i,j}(\sigma)\mathrm{d}\sigma = 0 \tag{14}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin(\omega_i\sigma)\sin(\omega_j\sigma)N_{i,i}(\sigma)\mathrm{d}\sigma = 0, \tag{15}$$

for all distinct i and j. Noting that Π is the common period of the probing frequencies we have

$$\frac{1}{\Pi}\int_0^{\Pi}\sin^2(\omega_i\sigma)\cos(2\omega_i\sigma)\mathrm{d}\sigma = -\frac{1}{4} \tag{16}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin^2(\omega_i\sigma)\sin^2(\omega_j\sigma)\mathrm{d}\sigma = \frac{1}{4} \tag{17}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin^3(\omega_i)\sin(\omega_j)\mathrm{d}\sigma = 0 \tag{18}$$
$$\frac{1}{\Pi}\int_0^{\Pi}\sin(\omega_i\sigma)\sin(\omega_j\sigma)\cos(2\omega_i\sigma)\mathrm{d}\sigma = 0, \tag{19}$$

for all $i \neq j$. Hence, one possible choice of elements of the $n \times n$ matrix $N(t)$ that satisfy all of the aforementioned constraints is given by

$$N_{i,i} = \frac{16}{a_i^2}\left(\sin^2(\omega_i t) - \frac{1}{2}\right) \tag{20}$$
$$N_{i,j} = \frac{4}{a_i a_j}\sin(\omega_i t)\sin(\omega_j t), \quad i \neq j, \tag{21}$$

where $N^T(t) = N(t)$. Based on this selection, we have

$$\frac{1}{\Pi}\int_0^{\Pi}N(\sigma)y\mathrm{d}\sigma = H + \frac{1}{\Pi}\int_0^{\Pi}R(\tilde{\theta} + S(\sigma))N(\sigma)\mathrm{d}\sigma. \tag{22}$$

In Section 6 we show that this averaged value converges to the actual value of the Hessian, under specific conditions on ω_i and a_i.

Computing the inverse of the Hessian matrix is the next step. Calculating Γ, the estimate of the inverse of the Hessian, in an algebraic fashion creates difficulties when the matrix $\hat{H}$ is close to singularity, or it is indefinite. To deal with this problem, a dynamic estimator is employed to calculate the inverse of $\hat{H}$ using a Riccati equation. Consider the following filter

$$\dot{\mathcal{H}} = -\omega_r\mathcal{H} + \omega_r\hat{H}. \tag{23}$$

Note that the state of this filter converges to $\hat{H}$, an estimate of H. Denote $\Gamma = \mathcal{H}^{-1}$. Since $\dot{\Gamma} = -\Gamma\dot{\mathcal{H}}\Gamma$, then Eq. (23) is transformed to the differential Riccati equation

$$\dot{\Gamma} = \omega_r\Gamma - \omega_r\Gamma\hat{H}\Gamma. \tag{24}$$

The equilibria of the Riccati Eq. (24) are $\Gamma^* = 0_{n\times n}$ and $\Gamma^* = \hat{H}^{-1}$, provided $\hat{H}$ settles to a constant. Since $\omega_r > 0$, the equilibrium $\Gamma^* = 0$ is unstable, whereas the linearization of (24) around $\Gamma^* = \hat{H}^{-1}$ has the Jacobian $-\omega_r I$, so the equilibrium at $\Gamma^* = \hat{H}^{-1}$ locally exponentially stable. This shows that, after a transient, the Riccati equation converges to the actual value of the inverse of the Hessian matrix if $\hat{H}$ is a good estimate of H.

A good estimate of the region of attraction of the exponentially stable equilibrium $\Gamma^* = \hat{H}^{-1}$ of (24) is difficult to obtain. An easy but conservative estimate makes the region of attraction inversely proportional to the largest eigenvalue of $\hat{H}$, which, due to the convergence of $\hat{H}$ to H, which we shall prove to be achieved locally (in an average sense), means that an estimate of the region of attraction of $\Gamma^* = H^{-1}$ is $1/\lambda_{\max}\{H\}$.

Remark 2. To highlight the contrast between the Newton and gradient algorithms, we refer to Remark 1 where the average behavior of the gradient algorithm is discussed. For the Newton algorithm in Fig. 2, the algorithm is given by

$$\dot{\hat{\theta}} = -K\Gamma M(t)y \tag{25}$$
$$\dot{\Gamma} = \omega_r\Gamma - \omega_r\Gamma N(t)y\Gamma, \tag{26}$$

where $\theta = \hat{\theta} + S(t)$. In the error variables $\tilde{\theta} = \hat{\theta} - \theta^*$, $\tilde{\Gamma} = \Gamma - H^{-1}$, when the map is quadratic, $Q(\theta) = Q^* + \frac{1}{2}(\theta-\theta^*)^T H(\theta-\theta^*)$, the averaged closed-loop system is given by

$$\dot{\tilde{\theta}} = -K\tilde{\theta} - K\tilde{\Gamma}H\tilde{\theta} \tag{27}$$
$$\dot{\tilde{\Gamma}} = -\omega_r\tilde{\Gamma} - \omega_r\tilde{\Gamma}H\tilde{\Gamma}, \tag{28}$$

where $K\tilde{\Gamma}H\tilde{\theta}$ is quadratic in $(\tilde{\Gamma}, \tilde{\theta})$, and $\omega_r\tilde{\Gamma}H\tilde{\Gamma}$ is quadratic in $\tilde{\Gamma}$. The linearization of this system has all of its eigenvalues at $-K$ and $-\omega_r$. Hence, unlike the gradient algorithm, whose convergence is governed by the unknown Hessian H, the convergence rate of the Newton algorithm can be arbitrarily assigned by the designer with an appropriate choice of K and ω_r.

4. Newton algorithm for dynamic systems

Consider a general multi-input-single-output (MISO) nonlinear model

$$\dot{x} = f(x, u), \tag{29}$$
$$y = h(x), \tag{30}$$

where $x \in \mathbb{R}^m$ is the state, $u \in \mathbb{R}^n$ is the input, $y \in \mathbb{R}$ is the output, and $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^m \to \mathbb{R}$ are smooth. Suppose that we know a smooth control law $u = \alpha(x, \theta)$ parameterized by a vector parameter $\theta \in \mathbb{R}^n$. The closed loop system $\dot{x} = f(x, \alpha(x, \theta))$ then has equilibria parameterized by θ. We make the following assumptions about the closed-loop system, as in Krstić and Wang (2000).

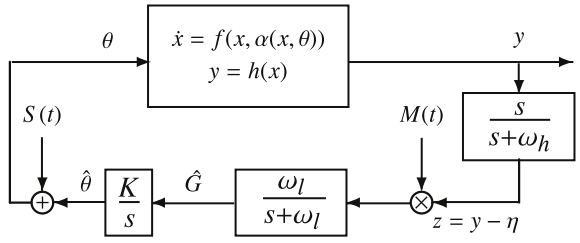

Fig. 3. Gradient-based extremum seeking.

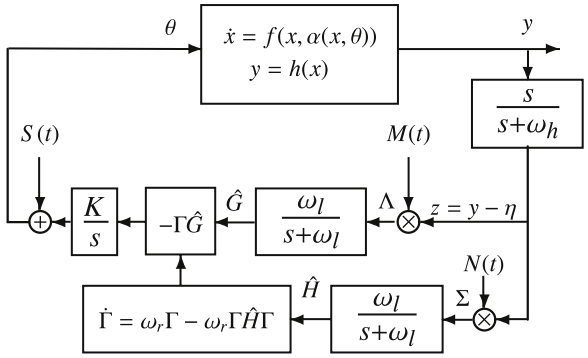

Fig. 4. Newton-based extremum seeking. The initial condition $\Gamma(0)$ should be chosen negative definite and symmetric.

Assumption 1. There exists a smooth function $l : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(x, \alpha(x, \theta)) = 0$ if and only if $x = l(\theta)$.

Assumption 2. For each $\theta \in \mathbb{R}^n$, the equilibrium $x = l(\theta)$ of the system $\dot{x} = f(x, \alpha(x, \theta))$ is locally exponentially stable uniformly in θ.

Assumption 3. There exists $\theta^* \in \mathbb{R}^n$ such that

$$\frac{\partial}{\partial\theta}(h \circ l)(\theta^*) = 0, \tag{31}$$

$$\frac{\partial^2}{\partial\theta^2}(h \circ l)(\theta^*) = H < 0, \quad H = H^T. \tag{32}$$

Our objective is to develop a feedback mechanism which maximizes the steady-state value of y but without requiring the knowledge of either θ^* or the functions h and l. The gradient-based extremum seeking design that achieves this objective, suitably adapted from Krstić and Wang (2000) to the multivariable case, is shown schematically in Fig. 3. Parallel to this, we present the generalized scheme for multivariable Newton-based extremum seeking as shown in Fig. 4.

The perturbation signals are defined by Eqs. (2), (3), (20) and (21). The probing frequencies ω_i's, the filter coefficients ω_h, ω_l and ω_r and gain K are selected as

$$\omega_i = \omega\omega_i' = O(\omega), \quad i \in \{1, 2, \ldots, n\} \tag{33}$$

$$\omega_h = \omega\omega_H = \omega\delta\omega_H' = O(\omega\delta) \tag{34}$$

$$\omega_l = \omega\omega_L = \omega\delta\omega_L' = O(\omega\delta) \tag{35}$$

$$\omega_r = \omega\omega_R = \omega\delta\omega_R' = O(\omega\delta) \tag{36}$$

$$K = \omega K' = \omega\delta K'' = O(\omega\delta), \tag{37}$$

where ω and δ are small positive constants, ω_i' is a rational number, ω_H', ω_L', and ω_R' are $O(1)$ positive constants, K'' is a $n \times n$ diagonal matrix with $O(1)$ positive elements, and $K' = \delta K''$.

The analysis of Ariyur and Krstić (2002), Krstić and Wang (2000) and Rotea (2000) shows that, in the gradient-based scheme, for "sufficiently small" ω and $|a|$, where $a = [a_1\ a_2\ \cdots\ a_n]^T$, and sufficiently small δ, which imply small filter cut-off frequencies, the states $(x, \hat{\theta})$ of the closed-loop system exponentially converge to an $O(\omega + \delta + |a|)$-neighborhood of $(l(\theta^*), \theta^*)$, and the output y converges to an $O(\omega + \delta + |a|)$-neighborhood of the optimum output $y^* = (h \circ l)(\theta^*)$.

In Section 6 we prove that the average value of $\Sigma(t)$ over the period Π is close enough to the actual value of the Hessian, under specific conditions on ω, δ and a. Since we are integrating over a finite time period, and we set the phase delays of the periodic perturbation signals equal to zero, it is possible to exclude condition $\omega_i \neq \omega_j + \omega_k$. The probing frequencies need to satisfy

$$\omega_i' \notin \left\{\omega_j', \frac{1}{2}(\omega_j' + \omega_k'), \omega_j' + 2\omega_k', \omega_j' + \omega_k' \pm \omega_l'\right\}, \tag{38}$$

for all distinct i, j, k, and l. As we see in Section 6, ignoring the conditions (38) is shifting the estimate of the parameter away from its true value, and leading to inaccurate estimates of the gradient vector and Hessian matrix.

5. Stability of the closed loop system with the Newton-based extremum seeking algorithm

We summarize the system in Fig. 4 as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x \\ \tilde{\theta} \\ \hat{G} \\ \tilde{\Gamma} \\ \tilde{H} \\ \tilde{\eta} \end{bmatrix} = \begin{bmatrix} f(x, \alpha(x, \theta^* + \tilde{\theta} + S(t))) \\ -K(\tilde{\Gamma} + H^{-1})\hat{G} \\ -\omega_l\hat{G} + \omega_l(y - h \circ l(\theta^*) - \tilde{\eta})M(t) \\ \omega_r(\tilde{\Gamma} + H^{-1})(I - (\tilde{H} + H)(\tilde{\Gamma} + H^{-1})) \\ -\omega_l\tilde{H} - \omega_l H + \omega_l(y - h \circ l(\theta^*) - \tilde{\eta})N(t) \\ -\omega_h\tilde{\eta} + \omega_h(y - h \circ l(\theta^*)) \end{bmatrix}. \tag{39}$$

To conduct a stability analysis we have introduced error variables $\tilde{\theta} = \hat{\theta} - \theta^*$, $\theta = \hat{\theta} + S(t)$, $\tilde{\eta} = \eta - h \circ l(\theta^*)$, $\tilde{\Gamma} = \Gamma - H^{-1}$, and $\tilde{H} = \hat{H} - H$, where η is governed by

$$\dot{\eta} = -\omega_h\eta + \omega_h y. \tag{40}$$

We perform a slight abuse of notation by stacking matrix quantities $\tilde{\Gamma}$ and $\tilde{H}$ along with vector quantities, as alternative notational choices would be more cumbersome.

Our main stability result is stated in the following theorem.

Theorem 1. *Consider the feedback system (39) under Assumptions 1–3. There exists $\bar{\omega} > 0$ and for any $\omega \in (0, \bar{\omega})$ there exist $\bar{\delta}, \bar{a} > 0$ such that for the given ω and any $|a| \in (0, \bar{a})$ and $\delta \in (0, \bar{\delta})$ there exists a neighborhood of the point $(x, \hat{\theta}, \hat{G}, \Gamma, \hat{H}, \eta) = (l(\theta^*), \theta^*, 0, H^{-1}, H, h \circ l(\theta^*))$ such that any solution of systems (39) from the neighborhood exponentially converges to an $O(\omega + \delta + |a|)$-neighborhood of that point. Furthermore, $y(t)$ converges to an $O(\omega + \delta + |a|)$-neighborhood of $h \circ l(\theta^*)$.*

To prepare for the proof of Theorem 1, which is given in Sections 6 and 7, and the Appendix, we summarize the system (39) in the time scale $\tau = \omega t$ as

$$\omega\frac{\mathrm{d}x}{\mathrm{d}\tau} = f(x, \alpha(x, \theta^* + \tilde{\theta} + \bar{S}(\tau))) \tag{41}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix} \tilde{\theta} \\ \hat{G} \\ \tilde{\Gamma} \\ \tilde{H} \\ \tilde{\eta} \end{bmatrix} = \delta\begin{bmatrix} -K''(\tilde{\Gamma} + H^{-1})\hat{G} \\ -\omega_L'\hat{G} + \omega_L'(y - h \circ l(\theta^*) - \tilde{\eta})\bar{M}(\tau) \\ \omega_R'(\tilde{\Gamma} + H^{-1})(I - (\tilde{H} + H)(\tilde{\Gamma} + H^{-1})) \\ -\omega_L'(\tilde{H} + H) + \omega_L'(y - h \circ l(\theta^*) - \tilde{\eta})\bar{N}(\tau) \\ -\omega_H'\tilde{\eta} + \omega_H'(y - h \circ l(\theta^*)) \end{bmatrix}, \tag{42}$$

where $\bar{S}(\tau) = S(t/\omega)$, $\bar{M}(\tau) = M(t/\omega)$ and $\bar{N}(\tau) = N(t/\omega)$.

6. Averaging analysis

The first step in our analysis is to study the system in Fig. 4. We "freeze" x in (41) at its equilibrium value $x = l(\theta^* + \tilde{\theta} + \bar{S}(\tau))$ and substitute it into (42), getting the reduced system

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\tilde{\theta}_r\\ \hat{G}_r\\ \tilde{\Gamma}_r\\ \tilde{H}_r\\ \tilde{\eta}_r\end{bmatrix} = \delta\begin{bmatrix}-K''(\tilde{\Gamma}_r + H^{-1})\hat{G}_r\\ -\omega_L'\hat{G}_r + \omega_L'(\nu(\tilde{\theta}_r + \bar{S}(\tau)) - \tilde{\eta}_r)\bar{M}(\tau)\\ \omega_R'(\tilde{\Gamma}_r + H^{-1})(I + (\tilde{H}_r + H)(\tilde{\Gamma}_r + H^{-1}))\\ -\omega_L'\tilde{H}_r - \omega_L'H + \omega_L'(\nu(\tilde{\theta}_r + \bar{S}(\tau)) - \tilde{\eta}_r)\bar{N}(\tau)\\ -\omega_H'\tilde{\eta}_r + \omega_H'\nu(\tilde{\theta}_r + \bar{S}(\tau))\end{bmatrix}, \tag{43}$$

where $\nu(z) = h \circ l(\theta^* + z) - h \circ l(\theta^*)$. In view of Assumption 3, $\nu(0) = 0$, $\partial\nu(0)/\partial z = 0$, and $\partial^2\nu(0)/\partial z^2 = H < 0$.

To prove the overall stability of (39), first we show that the reduced system (43) has a unique exponentially stable periodic solution around its equilibrium.

Theorem 2. *Consider system (43) under Assumption 3. There exist $\bar{\delta}, \bar{a} > 0$ such that for all $\delta \in (0, \bar{\delta})$ and $|a| \in (0, \bar{a})$ system (43) has a unique exponentially stable periodic solution $(\tilde{\theta}_r^{\Pi}(\tau), \hat{G}_r^{\Pi}(\tau), \tilde{\Gamma}_r^{\Pi}(\tau), \tilde{H}_r^{\Pi}(\tau), \tilde{\eta}_r^{\Pi}(\tau))$ of period Π and this solution satisfies*

$$\left|\tilde{\theta}_{r,i}^{\Pi}(\tau) - \sum_{j=1}^{n} c_{j,j}^{i} a_j^2\right| \le O(\delta + |a|^3) \tag{44}$$

$$|\hat{G}_r^{\Pi}(\tau)| \le O(\delta) \tag{45}$$

$$\left|\tilde{\Gamma}_r^{\Pi}(\tau) + \sum_{i=1}^{n}\sum_{j=1}^{n} H^{-1}W^{i}H^{-1}c_{j,j}^{i}a_j^2\right| \le O(\delta + |a|^3) \tag{46}$$

$$\left|\tilde{H}_r^{\Pi}(\tau) - \sum_{i=1}^{n}\sum_{j=1}^{n} W^{i}c_{j,j}^{i}a_j^2\right| \le O(\delta + |a|^3) \tag{47}$$

$$\left|\tilde{\eta}_r^{\Pi}(\tau) - \frac{1}{4}\sum_{i=1}^{n} H_{i,i}a_i^2\right| \le O(\delta + |a|^4) \tag{48}$$

for all $\tau \ge 0$, where

$$\begin{bmatrix}c_{j,j}^{1}\\ \vdots\\ c_{j,j}^{i-1}\\ c_{j,j}^{i}\\ c_{j,j}^{i+1}\\ \vdots\\ c_{j,j}^{n}\end{bmatrix} = -\frac{1}{12}H^{-1}\begin{bmatrix}\frac{\partial^3\nu}{\partial z_j\partial z_1^2}(0)\\ \vdots\\ \frac{\partial^3\nu}{\partial z_j\partial z_{j-1}^2}(0)\\ \frac{3}{2}\frac{\partial^3\nu}{\partial z_j^3}(0)\\ \frac{\partial^3\nu}{\partial z_j\partial z_{j+1}^2}(0)\\ \vdots\\ \frac{\partial^3\nu}{\partial z_j\partial z_n^2}(0)\end{bmatrix}, \quad \forall i, j \in \{1, 2, \ldots, n\} \tag{49}$$

$$(W^{i})_{j,k} = \frac{\partial^3\nu(0)}{\partial z_i\partial z_j\partial z_k}, \quad \forall i, j, \text{ and } k \in \{1, 2, \ldots, n\}. \tag{50}$$

The proof is presented in the Appendix.

7. Singular perturbation analysis

Now, we address the full system in Fig. 4 whose state space model is given by (41) and (42) in the time scale $\tau = \omega t$. To make the notation in our further analysis compact, we write (42) as

$$\frac{\mathrm{d}\xi}{\mathrm{d}\tau} = \delta E(\tau, x, \xi), \tag{51}$$

where $\xi = (\tilde{\theta}, \hat{G}, \tilde{\Gamma}, \tilde{H}, \tilde{\eta})$. By Theorem 2, there exists an exponentially stable periodic solution $\xi_r^{\Pi}(\tau)$ such that

$$\frac{\mathrm{d}\xi_r^{\Pi}(\tau)}{\mathrm{d}\tau} = \delta E(\tau, L(\tau, \xi_r^{\Pi}(\tau)), \xi_r^{\Pi}(\tau)), \tag{52}$$

where $L(\tau, \xi) = l(\theta^* + \tilde{\theta} + \bar{S}(\tau))$. To bring the system (41) and (51) into the standard singular perturbation form, we shift the state ξ using the transformation $\tilde{\xi} = \xi - \xi_r^{\Pi}(\tau)$ and get

$$\frac{\mathrm{d}\tilde{\xi}}{\mathrm{d}\tau} = \delta\tilde{E}(\tau, x, \tilde{\xi}) \tag{53}$$

$$\omega\frac{\mathrm{d}x}{\mathrm{d}\tau} = \tilde{F}(\tau, x, \tilde{\xi}) \tag{54}$$

where

$$\tilde{E}(\tau, x, \tilde{\xi}) = E(\tau, x, \tilde{\xi} + \xi_r^{\Pi}(\tau)) - E(\tau, L(\tau, \xi_r^{\Pi}(\tau)), \xi_r^{\Pi}(\tau)) \tag{55}$$

$$\tilde{F}(\tau, x, \tilde{\xi}) = f\big(x, \alpha\big(\tilde{\xi}_1 + \theta^* + \tilde{\theta}_r^{\Pi}(\tau) + \bar{S}(\tau)\big)\big). \tag{56}$$

We note that $x = L(\tau, \tilde{\xi}_r + \xi_r^{\Pi}(\tau))$ is the quasi-steady state, and that the reduced model

$$\frac{\mathrm{d}\tilde{\xi}_r}{\mathrm{d}\tau} = \delta\tilde{E}(\tau, L(\tau, \tilde{\xi}_r + \xi_r^{\Pi}(\tau)), \tilde{\xi}_r + \xi_r^{\Pi}(\tau)) \tag{57}$$

has an equilibrium at the origin $\tilde{\xi}_r = 0$. This equilibrium has been shown in Section 6 to be exponentially stable for a small $|a|$.

To complete the singular perturbation analysis, we also study the boundary layer model (in the time scale $t - t_0 = \tau/\omega$):

$$\begin{aligned}\frac{\mathrm{d}x_{\mathrm{b}}}{\mathrm{d}t} &= \tilde{F}(\tau, x_{\mathrm{b}} + L(\tau, \tilde{\xi} + \xi_r^{\Pi}(\tau)), \tilde{\xi}),\\ &= f(x_{\mathrm{b}} + l(\theta), \alpha(x_{\mathrm{b}} + l(\theta), \theta)),\end{aligned} \tag{58}$$

where $\theta = \theta^* + \tilde{\theta} + \bar{S}(\tau)$ should be viewed as a parameter independent from the time variable t. Since $f(l(\theta), \alpha(l(\theta), \theta)) \equiv 0$, then $x_{\mathrm{b}} \equiv 0$ is an equilibrium of (58). By Assumption 2, this equilibrium is locally exponentially stable uniformly in θ (and hence $l(\theta)$).

By combining exponential stability of the reduced model (57) with the exponential stability of the boundary layer model (58), using Tikhonov's theorem on the Infinite Interval (Theorem 9.4 in Khalil (1996)), we conclude the following:

(a) The solution $\xi(\tau)$ of (51) is $O(\omega)$-close to the solution $\xi_r(\tau)$ of (57), and therefore, it exponentially converges to an $O(\omega)$-neighborhood of the periodic solution $\xi_r^{\Pi}(\tau)$, which is $O(\delta)$-close to the equilibrium $\xi_r^{a,e}$. This, in turn, implies that the solution $\tilde{\theta}(\tau)$ of (42) exponentially converges to an $O(\omega + \delta)$-neighborhood of

$$\sum_{j=1}^{n}[c_{j,j}^{1}c_{j,j}^{2}\cdots c_{j,j}^{n}]^{T}a_j^2 + [O(|a|^3)]_{n\times 1}. \tag{59}$$

It follows then that $\theta(\tau) = \theta^* + \tilde{\theta}(\tau) + \bar{S}(\tau)$ exponentially converges to an $O(\omega + \delta + |a|)$-neighborhood of θ^*.

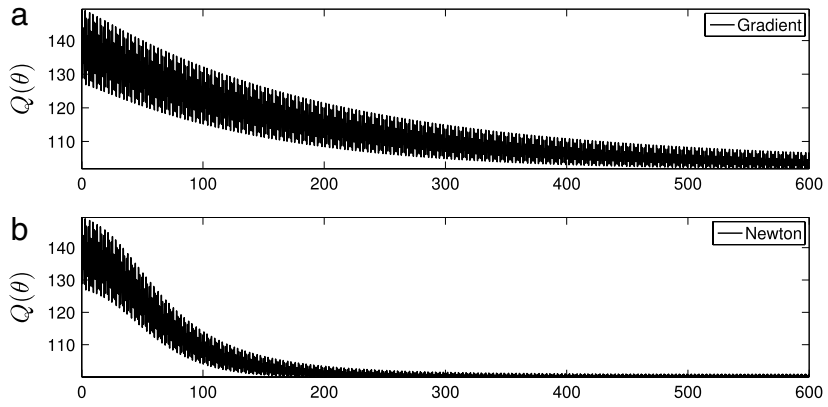

Fig. 5. The estimate of the maximum versus time.

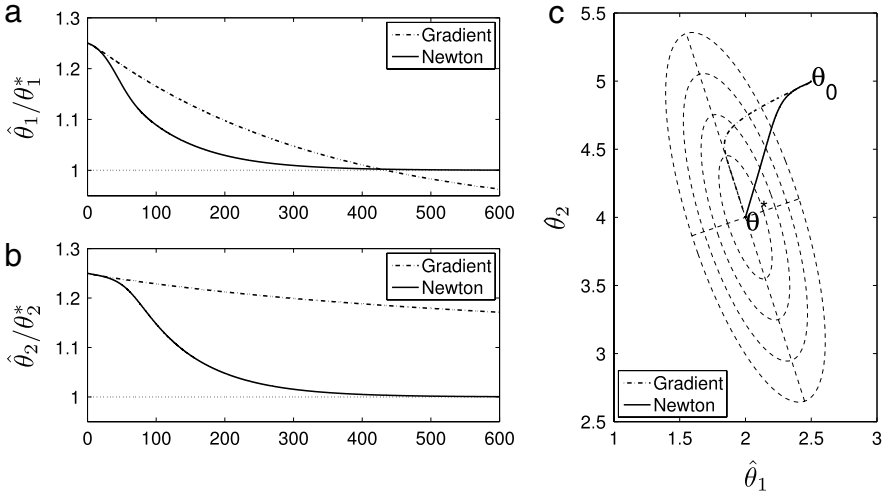

Fig. 6. Parameter estimates. (a) and (b) Time responses. (c) Phase portrait. The Newton trajectory is straight to the extremum, whereas the gradient trajectory follows the curved, steepest-descent path.

(b) The solution $x(\tau)$ of (54) satisfies

$$x(\tau) - l(\theta^* + \tilde{\theta}_r(\tau) + \bar{S}(\tau)) - x_b(t) = O(\omega), \tag{60}$$

where $\tilde{\theta}_r(\tau)$ is the solution of the reduced model (43) and $x_b(t)$ is the solution of the boundary layer model (58). From (60) we get

$$x(\tau) - l(\theta^*) = O(\omega) + l(\theta^* + \tilde{\theta}_r(\tau) + \bar{S}(t)) - l(\theta^*) + x_b(t). \tag{61}$$

Since $\tilde{\theta}_r(\tau)$ exponentially converges to the periodic solution $\tilde{\theta}_r^{\Pi}(\tau)$, which is $O(\delta)$-close to the average equilibrium (59), and since the solution $x_b(t)$ of (58) is exponentially decaying, then by (61), $x(\tau) - l(\theta^*)$ exponentially converges to an $O(\omega + \delta + |a|)$-neighborhood of zero. Consequently, $y = h(x)$ exponentially converges to an $O(\omega + \delta + |a|)$-neighborhood of its maximal equilibrium value $h \circ l(\theta^*)$.

This completes the proof of Theorem 1.

8. Simulation results

To illustrate the results and highlight the difference between the gradient-based and Newton-based extremum seeking methods, the following static quadratic input–output map is considered:

$$y = Q(\theta) = Q^* + \frac{1}{2}(\theta - \theta^*)^T H(\theta - \theta^*). \tag{62}$$

To make a fair comparison between the two methods, all parameters are chosen the same except the gain matrix. Before selecting the matrix K we investigate the performance of the gradient-based scheme versus the Newton-based scheme.

Recall (27) and (28). The initial convergence rate for the Newton-based scheme is governed by the time-varying matrix $-K_n\Gamma(t)H$. Eq. (4) shows that in the gradient-based scheme the convergence depends on the eigenvalues of K_gH. This means that, to have a fair comparison between the two methods, we should select K_g and K_n such that $K_g = -K_n\Gamma(0)$.

We perform our tests with the following parameters, $\delta = 0.1$, $\omega = 0.1$ rad/s, $\omega_1 = 70\omega$, $\omega_2 = 50\omega$, $\omega_L' = 10$, $\omega_H' = 8$, $\omega_R' = 10$, $a = [0.1\ 0.1]^T$, $K_g'' = 10^{-4}\,\mathrm{diag}([-25\ -25])$, $K_n'' = \mathrm{diag}([1\ 1])$, $\Gamma_0^{-1} = 400\,\mathrm{diag}([11])$, $\hat{\theta}_0 = [2.5\ 5]^T$, $Q^* = 100$, $\theta^* = [2\ 4]^T$, $H_{11} = 100$, $H_{12} = H_{21} = 30$, and $H_{22} = 20$.

Fig. 5 illustrates the estimate of the maximum. Evolution of the parameters is depicted in Fig. 6. Since the initial estimate of the Hessian is not true, each parameter starts to update with a different rate. As seen in Fig. 7, after 40 s the estimate of the Hessian is close enough to its actual value. Hence, the convergence rates of both parameters are the same after 40 s. Furthermore, Fig. 6(c) shows that, except for a short initial transient that is due to the estimation of the inverse of the Hessian, the Newton-based extremum seeking moves the parameters to the peak along a straight trajectory. In contrast, the trajectory of the gradient-based algorithm is curved and of greater length. Fig. 6(a) and (c) show that the gradient algorithm, which follows the steepest-descent path, results in the parameter $\hat{\theta}_1$ undershooting below its true value, which is not the case with the Newton algorithm whose parameter transients are monotonic. The Hessian matrix converges to its actual value as depicted in Fig. 7. Also it is worth noting that the Hessian converges

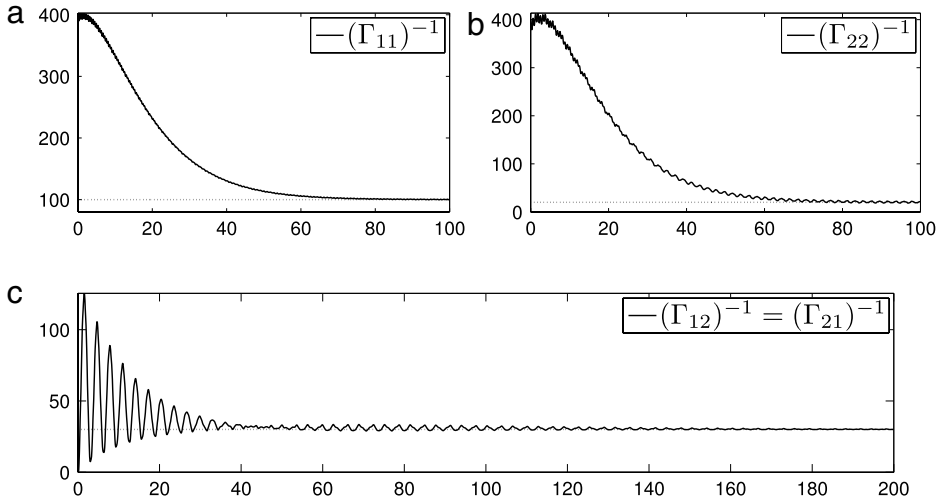

Fig. 7. Time evolution of the Hessian matrix estimator Γ^{-1}. The true value of H is reached in 40 s. Note in Fig. 6 that the Newton and gradient trajectories coincide for the first 40 s, after which Newton takes a straight path.

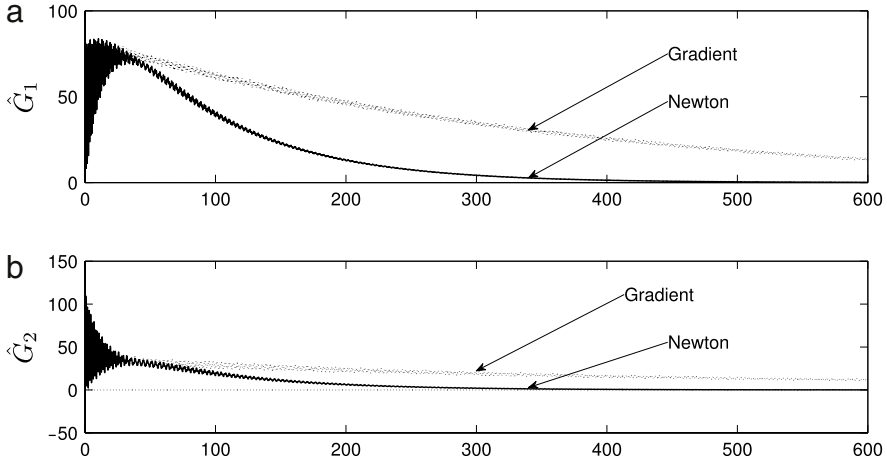

Fig. 8. The estimate of the gradient vector versus time.

faster to its actual value than $\hat{G}$ and $\hat{\theta}$. As illustrated in Fig. 8 the estimate of the gradient vector converges to zero after the Hessian matrix finds its true value.

9. Conclusions

Using the gradient-based extremum seeking in the multivariable case without having good information about the curvature of the cost function, namely, the Hessian matrix, may result inappropriate performance. With a growing number of the parameters, it is almost impossible to tune the convergence rate of all parameters in a desirable fashion. The Newton-based extremum seeking, which relies on the estimation of the gradient and Hessian matrix of the cost function at the same time, removes the trial and error process to update all parameters uniformly. Furthermore, the proposed Newton scheme eliminates the concern about the inversion of the Hessian estimate matrix by performing the inversion dynamically using a Riccati equation filter. The convergence rates of both the parameter and of the estimator of the Hessian inverse are independent of the unknown Hessian and can be assigned arbitrary by the user.

Appendix. Proof of Theorem 2

System (43) is in the form to which the averaging method is applicable. The average model of (43) is

$$\frac{\mathrm{d}}{\mathrm{d}\tau}[\tilde{\theta}_r^{aT} \quad \hat{G}_r^{aT} \quad \tilde{\Gamma}_r^{aT} \quad \tilde{H}_r^{aT} \quad \tilde{\eta}_r^{aT}]^T = \delta \begin{bmatrix} -K''(\tilde{\Gamma}_r^a + H^{-1})\hat{G}_r^a \\ -\omega_L'\hat{G}_r^a + \omega_L'\dfrac{1}{\Pi}\displaystyle\int_0^{\Pi} \nu(\tilde{\theta}_r^a + \bar{S}(\sigma))\bar{M}(\sigma)\mathrm{d}\sigma \\ \omega_R'(\tilde{\Gamma}_r^a + H^{-1})(I - (\tilde{H}_r^a + H)(\tilde{\Gamma}_r^a + H^{-1})) \\ -\omega_L'\tilde{H}_r^a - \omega_L'H + \omega_L'\dfrac{1}{\Pi}\displaystyle\int_0^{\Pi} \nu(\tilde{\theta}_r^a + \bar{S}(\sigma))\bar{N}(\sigma)\mathrm{d}\sigma \\ -\omega_H'\tilde{\eta}_r^a + \omega_H'\dfrac{1}{\Pi}\displaystyle\int_0^{\Pi} \nu(\tilde{\theta}_r^a + \bar{S}(\sigma))\mathrm{d}\sigma \end{bmatrix}. \tag{63}$$

The average equilibrium $(\tilde{\theta}_r^{a,e}, \hat{G}_r^{a,e}, \tilde{\Gamma}_r^{a,e}, \tilde{H}_r^{a,e}, \tilde{\eta}_r^{a,e})$ satisfies

$$\hat{G}_r^{a,e} = 0 \tag{64}$$

$$\int_0^{\Pi} \nu(\tilde{\theta}_r^{a,e} + \bar{S}(\sigma))\bar{M}(\sigma)\mathrm{d}\sigma = 0 \tag{65}$$

$$\tilde{\eta}_r^{a,e} = \frac{1}{\Pi}\int_0^{\Pi} \nu(\tilde{\theta}_r^{a,e} + \bar{S}(\sigma))\mathrm{d}\sigma \tag{66}$$

$$\tilde{H}_r^{a,e} + H = \frac{1}{\Pi}\int_0^{\Pi} \nu(\tilde{\theta}_r^{a,e} + \bar{S}(\sigma))\bar{N}(\sigma)\mathrm{d}\sigma \tag{67}$$

$$(\tilde{H}_r^{a,e} + H)(\tilde{\Gamma}_r^{a,e} + H^{-1}) = I. \tag{68}$$

As we shall see, for small $|a|$, $\tilde{\Gamma}_r^{a,e} + H^{-1} < 0$. By postulating $\tilde{\theta}_{r,i}^{a,e}$ in the form

$$\tilde{\theta}_{r,i}^{a,e} = \sum_{j=1}^{n} b_j^i a_j + \sum_{j=1}^{n}\sum_{k=1}^{n} c_{j,k}^i a_j a_k + O(|a|^3), \tag{69}$$

where b_j^i and $c_{j,k}^i$ are real numbers, substituting (69) in (64)–(68) and equating the like powers of a_j, we get $b_j^i = 0$, for all $i, j \in \{1, 2, \ldots, n\}$, $c_{j,k}^i = 0$, for all $i, j \neq k \in \{1, 2, \ldots, n\}$, and (49). The equilibrium of the average system is:

$$\begin{aligned}
\tilde{\theta}_{r,i}^{a,e} &= \sum_{j=1}^{n} c_{j,j}^i a_j^2 + O(|a|^3)\\
\hat{G}_r^{a,e} &= 0_{n\times 1}\\
\tilde{\Gamma}_r^{a,e} &= -\sum_{i=1}^{n}\sum_{j=1}^{n} H^{-1} W^i H^{-1} c_{j,j}^i a_j^2 + [O(|a|^3)]_{n\times n}\\
\tilde{H}_r^{a,e} &= \sum_{i=1}^{n}\sum_{j=1}^{n} W^i c_{j,j}^i a_j^2 + [O(|a|^3)]_{n\times n}\\
\tilde{\eta}_r^{a,e} &= \frac{1}{4}\sum_{i=1}^{n} H_{i,i} a_i^2 + O(|a|^4),
\end{aligned} \tag{70}$$

where W^i is a $n \times n$ matrix defined by (50). The Jacobian of the average system at equilibrium is

$$J_r^{a,e} = \delta \begin{bmatrix} A_{2n\times 2n} & 0_{2n\times(2n+1)} \\ B_{(2n+1)\times 2n} & C_{(2n+1)\times(2n+1)} \end{bmatrix}, \tag{71}$$

$$A = \begin{bmatrix} 0_{n\times n} & -K''(H^{-1} + \tilde{\Gamma}_r^{a,e}) \\ \dfrac{\omega_L'}{\Pi}\displaystyle\int_0^{\Pi} \frac{\partial}{\partial\tilde{\theta}}(\nu\bar{M})\mathrm{d}\sigma & -\omega_L' I_{n\times n} \end{bmatrix},$$

$$B = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} \\ \omega_L'\dfrac{1}{\Pi}\displaystyle\int_0^{\Pi} \frac{\partial}{\partial\tilde{\theta}}(\nu\bar{N})\mathrm{d}\sigma & 0_{n\times n} \\ \omega_H'\dfrac{1}{\Pi}\displaystyle\int_0^{\Pi} \frac{\partial}{\partial\tilde{\theta}}(\nu)\mathrm{d}\sigma & 0_{1\times n} \end{bmatrix},$$

$$C = \begin{bmatrix} -\omega_R' I_{n\times n} + \mathcal{O}_1 & -\omega_R' H^{-2} + \mathcal{O}_2 & 0_{n\times 1} \\ 0_{n\times n} & -\omega_L' I_{n\times n} & 0_{n\times 1} \\ 0_{1\times n} & 0_{1\times n} & -\omega_H' \end{bmatrix},$$

$$\mathcal{O}_1 = \omega_R' \sum_{i=1}^{n}\sum_{j=1}^{n} H^{-1} W^i c_{j,j}^i a_j^2 + [O(|a|^3)],$$

$$\begin{aligned}
\mathcal{O}_2 \;=\; & \omega_R' \sum_{i=1}^{n}\sum_{j=1}^{n} H^{-1}(W^i H^{-1} - H^{-1} W^i) H^{-1} c_{j,j}^i a_j^2 \\
& + [O(|a|^3)].
\end{aligned}$$

Since $J_r^{a,e}$ is block-lower-triangular, it is Hurwitz if and only if

$$A_{21} := \frac{\omega_L'}{\Pi}\int_0^{\Pi} \bar{M}(\sigma)\frac{\partial}{\partial\tilde{\theta}}\nu(\tilde{\theta}_r^{a,e} + \bar{S}(\sigma))\mathrm{d}\sigma < 0. \tag{72}$$

With a Taylor expansion we get that $A_{21} = \omega_L' H + O(|a|)$. We then have

$$\begin{aligned}
&\det(\lambda I_{2n\times 2n} - \delta A)\\
&\quad = \det(\lambda(\lambda + \omega_L'\delta) I_{n\times n} + \delta^2 K''(H^{-1} + \tilde{\Gamma}_r^{a,e})\,(A_{21}))\\
&\quad = \det((\lambda^2 + \omega_L'\delta\lambda) I_{n\times n}\\
&\qquad + \delta^2 K''(H^{-1} + [O(|a|^2)])(\omega_L' H + [O(|a|)]))\\
&\quad = \det\left((\lambda^2 + \omega_L'\delta\lambda) I_{n\times n} + \omega_L'\delta^2 K'' + [O(\delta^2|a|)]_{n\times n}\right),
\end{aligned} \tag{73}$$

which, in view of $H < 0$, proves that $J_r^{a,e}$ is Hurwitz for a that is sufficiently small in norm. This implies that the equilibrium (70) of the average system (63) is exponentially stable if all elements of vector a are sufficiently small. Then, according to the averaging theorem (Khalil, 1996), the proof is completed.

References

Airapetyan, R. (1999). Continuous newton method and its modification. *Applicable Analysis*, *73*, 463–484.

Ariyur, K.B., & Krstić, M. (2002). Analysis and design of multivariable extremum seeking. In *Proc. of the American control conference*.

Ariyur, K. B., & Krstić, M. (2003). *Real-time optimization by extremum seeking feedback*. Wiley-Interscience.

Banaszuk, A., Ariyur, K. B., Krstić, M., & Jacobson, C. A. (2004). An adaptive algorithm for control of combustion instability. *Automatica*, *40*, 1965–1972.

Becker, R., King, R., Petz, R., & Nitsche, W. (2007). Adaptive closed-loop separation control on a high-lift configuration using extremum seeking. *AIAA Journal*, *45*, 1382–1392.

Carnevale, D., Astolfi, A., Centioli, C., Podda, S., Vitale, V., & Zaccarian, L. (2009). A new extremum seeking technique and its application to maximize RF heating on FTU. *Fusion Engineering and Design*, *84*, 554–558.

Choi, J.-Y., Krstić, M., Ariyur, K. B., & Lee, J. S. (2002). Extremum seeking control for discrete-time systems. *IEEE Transactions on Automatic Control*, *47*, 318–323.

Cochran, J., Kanso, E., Kelly, S. D., Xiong, H., & Krstić, M. (2009). Source seeking for two nonholonomic models of fish locomotion. *IEEE Transactions on Robotics Automation*, *25*, 1166–1176.

Cochran, J., & Krstić, M. (2009). Nonholonomic source seeking with tuning of angular velocity. *IEEE Transactions on Automatic Control*, *54*, 717–731.

Guay, M., Perrier, M., & Dochain, D. (2005). Adaptive extremum seeking control of nonisothermal continuous stirred reactors. *Chemical Engineering Science*, *60*, 3671–3681.

Khalil, H. K. (1996). *Nonlinear systems* (2nd ed.). NJ, Englewood Cliffs: Prentice Hall.

Killingsworth, N. J., Aceves, S. M., Flowers, D. L., Espinosa-Loza, F., & Krstić, M. (2009). HCCI engine combustion-timing control: optimizing gains and fuel consumption via extremum seeking. *IEEE Transactions on Control System Technology*, *17*, 1350–1361.

Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*, 595–601.

Luo, L., & Schuster, E. (2009). Mixing enhancement in 2d magnetohydrodynamic channel flow by extremum seeking boundary control. In *Proc. of the American control conference*.

Moase, W. H., Manzie, C., & Brear, M. J. (2010). Newton-like extremum-seeking for the control of thermoacoustic instability. *IEEE Transactions on Automatic Control*, *55*, 2094–2105.

Nešić, D., Tan, Y., Moase, W.H., & Manzie, C. (2010). A unifying approach to extremum seeking: adaptive schemes based on the estimation of derivatives. In *Proc. of IEEE conf. on decision and control*.

Rotea, M.A. Analysis of multivariable extremum seeking algorithms. In *Proc. of the American control conference*.

Stankovic, M.S., Johansson, K.H., & Stipanovic, D.M. (2010). Distributed seeking of nash equilibria in mobile sensor networks. In *Proc. of IEEE conf. on decision and control*.

Stankovic, M. S., & Stipanovic, D. M. (2010). Extremum seeking under stochastic noise and applications to mobile sensors. *Automatica*, *46*, 1243–1251.

Tan, Y., Nešić, D., & Mareels, I. (2006). On non-local stability properties of extremum seeking control. *Automatica*, *42*, 889–903.

Tee1, A.R., & Popovic, D. (2001). Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming. In *Proc. of the American control conference*.

Wang, H.-H., & Krstić, M. (2000). Extremum seeking for limit cycle minimization. *IEEE Transactions on Automatic Control*, *45*, 2432–2437.

Wang, H.-H., Yeung, S., & Krstić, M. (2000). Experimental application of extremum seeking on an axial-flow compressor. *IEEE Transactions on Control System Technology*, *8*, 300–309.

Zhang, C., Arnold, D., Ghods, N., Siranosian, A., & Krstić, M. (2007). Source seeking with nonholonomic unicycle without position measurement and with tuning of forward velocity. *Systems & Control Letters*, *56*, 245–252.

Azad Ghaffari received his B.S. degree in electrical engineering and his M.S. degree in control engineering from K.N. Toosi University of Technology in Tehran, Iran. He is now working towards his Ph.D. degree in the Joint Doctoral Program in Mechanical and Aerospace Engineering at the University of California, San Diego, and San Diego State University. His research interests include extremum seeking, control of photovoltaic arrays, power electronics, and sliding mode control.

Miroslav Krstic is the Daniel L. Alspach Professor and the founding Director of the Cymer Center for Control Systems and Dynamics (CCSD) at UC San Diego. He received his Ph.D. in 1994 from UC Santa Barbara and was Assistant Professor at the University of Maryland until 1997. He is a coauthor of eight books: Nonlinear and Adaptive Control Design (Wiley, 1995), Stabilization of Nonlinear Uncertain Systems (Springer, 1998), Flow Control by Feedback (Springer, 2002), Real-time Optimization by Extremum Seeking Control (Wiley, 2003), Control of Turbulent and Magnetohydrodynamic Channel Flows (Birkhauser, 2007), Boundary Control of PDEs: A Course on Backstepping Designs (SIAM, 2008), Delay Compensation for Nonlinear, Adaptive, and PDE Systems (Birkhauser, 2009), and Adaptive Control of Parabolic PDEs (Princeton, 2010). Krstic is a Fellow of the IEEE and IFAC and has received the Axelby and Schuck paper prizes, NSF Career, ONR Young Investigator, and PECASE awards. He has held the appointment of Springer Distinguished Visiting Professor of Mechanical Engineering at UC Berkeley. He has served in senior editorial roles for Automatica, IEEE Transactions on Automatic Control, and Springer-Verlag.

Dragan Nešić received the B.E. degree in mechanical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1990, and the Ph.D. degree in systems engineering, RSISE, Australian National University, Canberra, Australia, in 1997. He is a Professor with the Electrical and Electronic Engineering Department, University of Melbourne, Melbourne, Australia. His research interests include networked control systems, discrete-time, sampled-data and continuous-time nonlinear control systems, input-to-state stability, extremum seeking control, applications of symbolic computation in control theory, hybrid control systems, and so on. Prof. Nešić is a Fellow of IEAust. He was a Distinguished Lecturer of the IEEE Control Systems Society from 2008 to 2010. He was awarded a Humboldt Research Fellowship in 2003 by the Alexander von Humboldt Foundation, and an Australian Professorial Fellowship (20042009) and Future Fellowship (20102014) by the Australian Research Council.